

VXLAN Data Plane Learning Lab (meccanismo di apprendimento in caso di alluvione)

Sommario

[Introduzione](#)

[Prerequisiti](#)

[Requisiti](#)

[Componenti usati](#)

[Premesse](#)

[Terminologie utilizzate](#)

[Configurazione](#)

[Esempio di rete](#)

[Configurazioni](#)

Introduzione

Questo documento descrive come configurare un CML con switch Nexus 9Kv usando la VXLAN con il metodo Flood and Learn.

Prerequisiti

Requisiti

Cisco raccomanda la conoscenza dei seguenti argomenti:

- Informazioni su routing e switching
- Nozioni base sul routing multicast, ad esempio Rendezvous Point (RP) e Platform Independent Multicast (PIM)

Componenti usati

Il documento può essere consultato per tutte le versioni software o hardware.

Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti. Se la rete è operativa, valutare attentamente eventuali conseguenze derivanti dall'uso dei comandi.

Premesse

Il documento fornisce anche linee guida per la distribuzione del lab, nonché per la verifica delle

configurazioni e delle operazioni.

In questo laboratorio, il Cisco Modeling Lab (CML) con switch Nexus 9000V viene utilizzato sia per Leaf che per Spine.

Foglia1	Loopback0 - 1.1.1.1	Loopback1 - 10.10.10.10
Foglia2	Loopback0 - 2.2.2.2	Loopback1 - 20.20.20.20
Foglia 3	Loopback0 - 3.3.3.3	Loopback1 - 30.30.30.30
Dorso 1	Loopback0 - 4.4.4.4	Loopback1 - 60.60.60.60 - Anycast RP
Dorso2	Loopback0 - 5.5.5.5	Loopback1 - 60.60.60.60 - Anycast RP
Subnet desktop	192.168.100.0/24	

Terminologie utilizzate

VTEP (Virtual eXtensible Local Area Network) Tunnel Endpoint (VTEP): incapsula il traffico MAC nel traffico IP e instrada il traffico MAC verso altri VTEP.

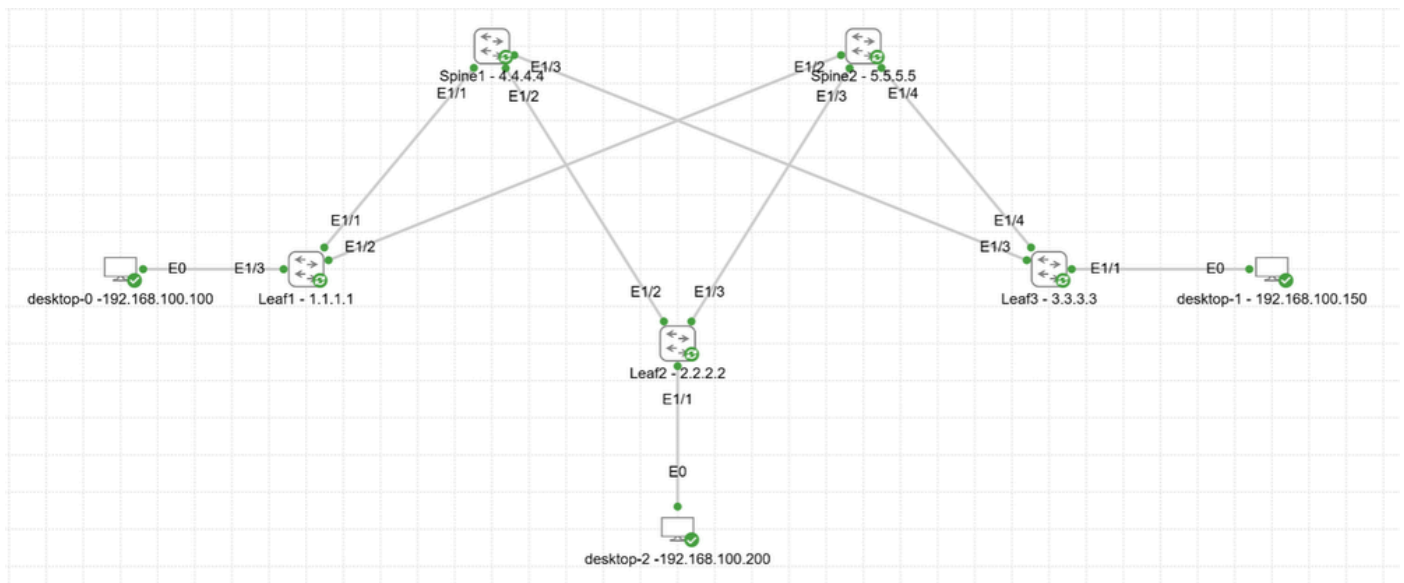
VXLAN Network Identifier (VNID): ID all'interno dell'intestazione della VXLAN che identifica la rete e può essere mappato a una VLAN. Da una prospettiva di inoltro, un VNID è un dominio di trasmissione.

Network Virtual Interface (NVE) - Interfaccia logica dove si verificano l'incapsulamento e la deincapsulamento.

Broadcast, unicast e multicast sconosciuti (BUM)

Configurazione

Esempio di rete



Esempio di connettività di rete

Configurazioni

Passaggio 1.

- Abilita la funzionalità OSPF (Open Shortest Path First).
- Aggiungere loopback a tutti i dispositivi.
- Abilitare OSPF su interfacce e loopback Ethernet.

```
leaf1# show running-config interface ethernet 1/1

!Command: show running-config interface Ethernet1/1
!Running configuration last done at: Tue Dec 24 13:12:55 2024
!Time: Wed Dec 25 05:24:23 2024

version 9.3(8) Bios:version

interface Ethernet1/1
  no switchport
  ip address 70.0.0.2/30
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

leaf1# show running-config interface loopback 0

!Command: show running-config interface loopback0
!Running configuration last done at: Tue Dec 24 13:12:55 2024
!Time: Wed Dec 25 05:24:24 2024

version 9.3(8) Bios:version

interface loopback0
  ip address 1.1.1.1/32
  ip router ospf 1 area 0.0.0.0
```

```

spinel# show running-config interface ethernet 1/1

!Command: show running-config interface Ethernet1/1
!Running configuration last done at: Tue Dec 24 13:16:16 2024
!Time: Wed Dec 25 05:25:46 2024

version 9.3(8) Bios:version

interface Ethernet1/1
  no switchport
  ip address 70.0.0.1/30
  ip router ospf 5 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

spinel# show running-config interface loopback 0

!Command: show running-config interface loopback0
!Running configuration last done at: Tue Dec 24 13:16:16 2024
!Time: Wed Dec 25 05:25:51 2024

version 9.3(8) Bios:version

interface loopback0
  ip address 4.4.4.4/32
  ip router ospf 5 area 0.0.0.0

```

Abilitazione di OSPF su interfaccia dorsale

Il vicinato OSPF è stabilito tra gli switch Leaf e Spine.

```

spinel# show ip ospf neighbors
OSPF Process ID 5 VRF default
Total number of neighbors: 3
Neighbor ID      Pri State           Up Time  Address      Interface
1.1.1.1          1 FULL/DR         16:22:51 70.0.0.2    Eth1/1
2.2.2.2          1 FULL/DR         16:22:52 50.0.0.2    Eth1/2
3.3.3.3          1 FULL/DR         16:22:52 30.0.0.2    Eth1/3

```

Protocollo adiacente OSPF stabilito con switch foglia

```

leaf1# show ip ospf neighbors
OSPF Process ID 1 VRF default
Total number of neighbors: 2
Neighbor ID      Pri State           Up Time  Address      Interface
4.4.4.4          1 FULL/BDR        16:15:40 70.0.0.1    Eth1/1
5.5.5.5          1 FULL/BDR        16:15:10 80.0.0.1    Eth1/2

```

Nodo adiacente OSPF stabilito con Spine Switch

```
leaf1# ping 3.3.3.3 source-interface loopback 0
PING 3.3.3.3 (3.3.3.3): 56 data bytes
64 bytes from 3.3.3.3: icmp_seq=0 ttl=253 time=6.616 ms
64 bytes from 3.3.3.3: icmp_seq=1 ttl=253 time=6.695 ms
64 bytes from 3.3.3.3: icmp_seq=2 ttl=253 time=6.018 ms
64 bytes from 3.3.3.3: icmp_seq=3 ttl=253 time=6.52 ms
64 bytes from 3.3.3.3: icmp_seq=4 ttl=253 time=7.768 ms
```

Ping raggiungibilità da Foglia1 a Foglia3

Passaggio 2.

Aggiungere un'interfaccia di loopback aggiuntiva che verrà utilizzata per la VXLAN sugli switch foglia. Verificare inoltre la raggiungibilità da tutti gli switch a foglia nel fabric.

```
leaf2# show running-config interface loopback 1

!Command: show running-config interface loopback1
!Running configuration last done at: Wed Dec 25 05:41:46 2024
!Time: Wed Dec 25 05:42:33 2024

version 9.3(8) Bios:version

interface loopback1
  description Vxlan loopback
  ip address 20.20.20.20/32
  ip router ospf 2 area 0.0.0.0
```

Loopback per VXLAN

```
leaf2# ping 10.10.10.10 source 20.20.20.20
PING 10.10.10.10 (10.10.10.10) from 20.20.20.20: 56 data bytes
64 bytes from 10.10.10.10: icmp_seq=0 ttl=253 time=7.187 ms
64 bytes from 10.10.10.10: icmp_seq=1 ttl=253 time=6.248 ms
64 bytes from 10.10.10.10: icmp_seq=2 ttl=253 time=5.472 ms
64 bytes from 10.10.10.10: icmp_seq=3 ttl=253 time=4.741 ms
64 bytes from 10.10.10.10: icmp_seq=4 ttl=253 time=4.887 ms

--- 10.10.10.10 ping statistics ---
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 4.741/5.707/7.187 ms
```

Raggiungibilità sovrapposizione da Foglia2 a Foglia1

Passaggio 3.

Configurare PIM Any-Source Multicast (ASM) e Anycast RP sugli aculei:

- Attivate la feature PIM.
- Attivate PIM su tutti i collegamenti sottostanti.
- Crea un nuovo loopback sugli aculei da utilizzare per Anycast RP.
- Annunciare questo loopback in OSPF.
- Configurare Anycast RP (funzione Nexus) sugli aculei.
- Configurare l'RP su tutti i dispositivi.

ANYCAST RP:

Anycast RP è un meccanismo che consente il rapido failover di RP e la condivisione del carico RP. L'RP Anycast implica l'utilizzo dello stesso indirizzo IP (indirizzo RP) su due o più router che funzioneranno come l'RP. Questo indirizzo IP deve essere annunciato nel protocollo IGP (Interior Gateway Protocol) in modo che altri router possano scegliere il percorso migliore per l'indirizzo IP. In caso di fallimento, il tempo di convergenza sarà lo stesso dell'IGP.

La presenza di più RP con lo stesso indirizzo IP garantisce che le origini e i riceventi vengano sempre indirizzati al RP più vicino in base alla tabella di routing unicast. I messaggi di aggiunta PIM dai ricevitori possono essere inviati a un RP, mentre i router designati da PIM registrano le loro origini locali a un altro RP.

È importante sincronizzare le informazioni tra i diversi RP in quanto alcuni mittenti e riceventi possono unirsi al router 1 come RP, mentre altri possono unirsi al router 2 come RP. Se i router non dispongono di informazioni complete su tutte le origini, la comunicazione multicast può essere interrotta. Per risolvere questo problema, è necessario un meccanismo che sincronizzi le informazioni sulle origini tra tutti i router che agiscono come server di inoltro della richiesta. A tale scopo, sono disponibili due protocolli: Protocollo MSDP (Multicast Source Discovery Protocol) e PIM.

```
spinel# show running-config interface loopback 1

!Command: show running-config interface loopback1
!Running configuration last done at: Tue Dec 24 13:16:16 2024
!Time: Wed Dec 25 05:50:44 2024

version 9.3(8) Bios:version

interface loopback1
  description Anycast RP loopback
  ip address 60.60.60.60/32
  ip router ospf 5 area 0.0.0.0
  ip pim sparse-mode

spinel# show running-config | section rp
ip pim rp-address 60.60.60.60 group-list 224.0.0.0/4
ip pim anycast-rp 60.60.60.60 4.4.4.4
ip pim anycast-rp 60.60.60.60 5.5.5.5
```

60.60.60.60 è l'indirizzo IP dell'indirizzo IP del destinatario della richiesta di offerta Anycast e 4.4.4.4/5.5.5.5 è l'indirizzo IP del

```
leaf2# show running-config interface ethernet 1/2

!Command: show running-config interface Ethernet1/2
!Running configuration last done at: Wed Dec 25 05:41:46 2024
!Time: Wed Dec 25 05:51:18 2024

version 9.3(8) Bios:version

interface Ethernet1/2
  no switchport
  ip address 50.0.0.2/30
  ip router ospf 2 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

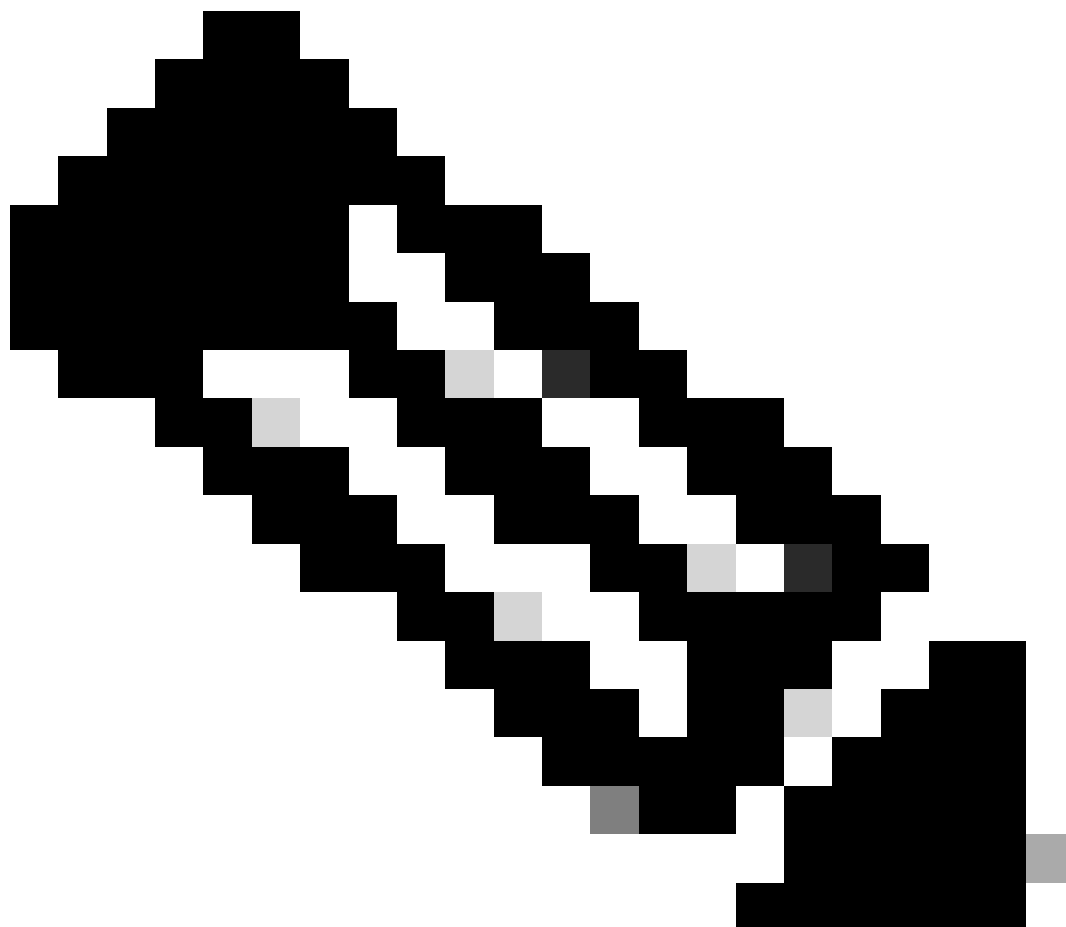
leaf2# show running-config | section rp-address
ip pim rp-address 60.60.60.60 group-list 224.0.0.0/4
```

Configurazione RP Anycast su switch foglia

```
spinel# show ip pim neighbor
PIM Neighbor Status for VRF "default"
Neighbor      Interface      Uptime      Expires      DR      Bidir-  BFD
ECMP Redirect                                     Priority Capable State
    Capable
70.0.0.2      Ethernet1/1    16:44:28    00:01:39    1        yes      n/a
no
50.0.0.2      Ethernet1/2    16:44:28    00:01:44    1        yes      n/a
no
30.0.0.2      Ethernet1/3    16:44:28    00:01:25    1        yes      n/a
no
spinel# show ip pim rp
PIM RP Status Information for VRF "default"
BSR disabled
Auto-RP disabled
BSR RP Candidate policy: None
BSR RP policy: None
Auto-RP Announce policy: None
Auto-RP Discovery policy: None

Anycast-RP 60.60.60.60 members:
  4.4.4.4*  5.5.5.5

RP: 60.60.60.60*, (0),
  uptime: 16:57:06  priority: 255,
  RP-source: (local),
  group ranges:
  224.0.0.0/4
```

Nota: Non dimenticare di posizionare il PIM sul loopback da usare per la VXLAN sugli switch foglia.

Passaggio 4.

- Abilitare la funzione VXLAN.
- Abilitare questa funzionalità per mappare le VLAN agli identificatori di rete virtuali (VNI).
- Crea NVE.
- Configurare la porta di accesso verso il desktop.

```
feature vn-segment-vlan-based
feature nv overlay
```

```
!Command: show running-config interface nve1
!Running configuration last done at: Wed Dec 25 06:08:01 2024
!Time: Wed Dec 25 06:08:04 2024

version 9.3(8) Bios:version

interface nve1
  no shutdown
  source-interface loopback1
  member vni 10000
  mcast-group 239.0.0.1
```

Creazione dell'interfaccia NVE

```
vlan 10
  vn-segment 10000
```

Mapping tra VLAN e segmento VN

```
leaf1# show interface nve 1
nve1 is up
admin state is up, Hardware: NVE
  MTU 9216 bytes
  Encapsulation VXLAN
  Auto-mdix is turned off
  RX
    ucast: 39 pkts, 3346 bytes - mcast: 0 pkts, 0 bytes
  TX
    ucast: 18 pkts, 2216 bytes - mcast: 0 pkts, 0 bytes
```

Stato dell'interfaccia NVE

Eseguire il ping tra Desktop0 e Desktop1 e Desktop2 per verificare che siano raggiungibili.

Quando viene avviata una richiesta ARP (Address Resolution Protocol) da Desktop0 a Desktop1, il pacchetto ARP viene inviato a Leaf1. Leaf1 inoltra quindi il pacchetto al dispositivo Spine utilizzando l'indirizzo multicast 239.0.0.1, utilizzato per VNI1000. Il dispositivo Spine esegue il multicast dei pacchetti a tutti i dispositivi foglia che fanno parte dello stesso VNI 10000.

Eseguire il ping tra Desktop0 (192.168.100.100), Desktop1 (192.168.100.150) e Desktop2 (192.168.100.200).

```

inserthostname-here:~$ ping 192.168.100.150
PING 192.168.100.150 (192.168.100.150): 56 data bytes
64 bytes from 192.168.100.150: seq=0 ttl=42 time=8.477 ms
64 bytes from 192.168.100.150: seq=1 ttl=42 time=12.791 ms
64 bytes from 192.168.100.150: seq=2 ttl=42 time=8.352 ms
^C
--- 192.168.100.150 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 8.352/9.873/12.791 ms
inserthostname-here:~$
inserthostname-here:~$
inserthostname-here:~$ ping 192.168.100.200
PING 192.168.100.200 (192.168.100.200): 56 data bytes
64 bytes from 192.168.100.200: seq=0 ttl=42 time=15.432 ms
64 bytes from 192.168.100.200: seq=1 ttl=42 time=9.228 ms
64 bytes from 192.168.100.200: seq=2 ttl=42 time=7.133 ms
^C
--- 192.168.100.200 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 7.133/10.597/15.432 ms
inserthostname-here:~$ arp -a
? (192.168.100.150) at 52:54:00:05:84:a2 [ether] on eth0
? (192.168.100.1) at 00:01:00:01:00:01 [ether] on eth0
? (192.168.100.200) at 52:54:00:10:70:ae [ether] on eth0

```

Eseguire il ping tra Desktop0 e Desktop1 e Desktop2

LEAF1 sta formando un peer NVE con LEAF3.

<#root>

```
leaf1# show nve peers
```

Interface	Peer-IP	State	LearnType	Uptime	Route
-----------	---------	-------	-----------	--------	-------

r-Mac

```

-----
-----

```

nve1

30.30.30.30

Up	DP	00:10:23	n/a
----	----	----------	-----

```
leaf1# show nve vni 10000
```

Codes: CP - Control Plane

DP - Data Plane

UC - Unconfigured

SA - Suppress ARP

- SU - Suppress Unknown Unicast
- Xconn - Crossconnect
- MS-IR - Multisite Ingress Replication

Interface	VNI	Multicast-group	State	Mode	Type	[BD/VRF]	Flags
nve1	10000	239.0.0.1	Up	DP	L2	[10]	

LEAF3 sta formando un peer NVE con LEAF1.

<#root>

leaf3# show nve peers

Interface	Peer-IP	State	LearnType	Uptime	Route
r-Mac					
nve1	10.10.10.10	Up	DP	00:10:56	n/a

leaf3# show nve vni 10000

- Codes: CP - Control Plane DP - Data Plane
- UC - Unconfigured SA - Suppress ARP
- SU - Suppress Unknown Unicast
- Xconn - Crossconnect
- MS-IR - Multisite Ingress Replication

Interface	VNI	Multicast-group	State	Mode	Type	[BD/VRF]	Flags
nve1	10000	239.0.0.1	Up	DP	L2	[10]	

leaf1# show mac address-table

Legend:

* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC

age - seconds since last seen,+ - primary entry using vPC Peer-Link,

(T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan

VLAN	MAC Address		Type	age	Secure	NTFY	Ports
-----+-----+-----+-----+-----+-----+-----							
* 10							
5254.0004.8b92							
dynamic	0	F	F	Eth1/3			----- MAC Address of De
* 10							
5254.0005.84a2							
dynamic	0	F	F	nve1(30.30.30.30)			----- MAC Address of Desktop1
G -	5206.ab8a.1b08	static	-	F	F	sup-eth1(R)	

leaf3# show mac address-table

Legend:

* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC

age - seconds since last seen,+ - primary entry using vPC Peer-Link,

(T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan

VLAN	MAC Address		Type	age	Secure	NTFY	Ports
-----+-----+-----+-----+-----+-----+-----							
* 10							
5254.0004.8b92							
dynamic	0	F	F	nve1(10.10.10.10)			----- MAC Address of Desktop0 con
* 10							
5254.0005.84a2							
dynamic	0	F	F	Eth1/1			----- MAC Address of Desk
G -	5206.0619.1b08	static	-	F	F	sup-eth1(R)	

Di seguito viene riportata l'istantanea di Wireshark quando un pacchetto ARP viene avviato da Leaf1 a un multicast.

No.	Time	Source	Destination	Protocol	Length	Info
7	5.105615	52:54:00:05:84:a2	52:54:00:04:8b:92	ARP	110	192.168.100.200 is at 52:54:00:05:84:a2
8	7.019252	52:54:00:04:8b:92	ff:ff:ff:ff:ff:ff	ARP	110	Who has 192.168.100.200? Tell 192.168.100.100

▶ Frame 8: 110 bytes on wire (880 bits), 110 bytes captured (880 bits)

▶ Ethernet II, Src: 52:06:ab:8a:1b:08 (52:06:ab:8a:1b:08), Dst: IPv4mcast_01 (01:00:5e:00:00:01)

▶ Internet Protocol Version 4, Src: 10.10.10.10, Dst: 239.0.0.1

▶ User Datagram Protocol, Src Port: 50384, Dst Port: 4789

▶ Virtual eXtensible Local Area Network

▶ Ethernet II, Src: RealtekU_04:8b:92 (52:54:00:04:8b:92), Dst: Broadcast (ff:ff:ff:ff:ff:ff)

▶ Address Resolution Protocol (request)

Wireshark Capture con il pacchetto di richiesta ARP inviato al gruppo multicast

Informazioni su questa traduzione

Cisco ha tradotto questo documento utilizzando una combinazione di tecnologie automatiche e umane per offrire ai nostri utenti in tutto il mondo contenuti di supporto nella propria lingua. Si noti che anche la migliore traduzione automatica non sarà mai accurata come quella fornita da un traduttore professionista. Cisco Systems, Inc. non si assume alcuna responsabilità per l'accuratezza di queste traduzioni e consiglia di consultare sempre il documento originale in inglese (disponibile al link fornito).